

AN APPLICATION OF RANK-ORDER STATISTICS TO THE JOINT SPATIAL AND TEMPORAL VARIATIONS OF METEOROLOGICAL ELEMENTS

PAUL R. JULIAN

Laboratory of Atmospheric Sciences, National Center for Atmospheric Research,* Boulder, Colo.

ABSTRACT

Principal components or empirical orthogonal functions are virtually the sole statistical tool used to date for investigations of space-time variability of meteorological elements. Maximum statistical efficiency and possible physical interpretation of empirical orthogonal functions derives from the assumptions of stationarity and homoscedasticity of the scalar variables in space and time. In this study, a rank table technique is given in which temporal data from a number of stations is ranked time-wise, and rank sums for each time obtained by summing ranks over the total number of stations. The technique offers some advantages for investigations of joint space-time variability. First, it is nonparametric; second, analysis of variance schemes are simplified; and third, a test of homoscedasticity can easily be performed. Networks of streamflow and precipitation data over the conterminous 48 States are used to illustrate the use of the technique. As a result, streamflow and precipitation data are shown to be spatially heteroscedastic—dry periods are better correlated spatially than wet periods. A runs test on the temporally varying rank sums suggests that while precipitation is not temporally heteroscedastic (dry and wet periods are both essentially randomly distributed), streamflow data might be. Apparently, years of deficient streamflow tend to be persistent while years of excessive streamflow are essentially randomly distributed in time.

1. INTRODUCTION

The joint space-time variability of meteorological elements seems to be a relatively unexplored subject. The most sophisticated method of analysis of such joint variability is the technique of empirical orthogonal functions or principal components (Lorenz 1956). Originating not because of intrinsic interest in space-time variations, but from the need for a statistically efficient and stable regression forecast technique, principal component analysis (PCA) has now been applied in a wide variety of problems. Examples are: parameterization of the vertical structure of pressure and temperature fields for numerical forecasting models by Holmström (1963), climatological representation of precipitation fields by Stidd (1967) and Sellers (1968), the representation of joint (multivariate) climatological fields by Kutzbach (1967), space-filtering of meteorological fields by Grimmer (1963), and for reproducing the relevant space-time properties of hydrometeorological variables in stochastic streamflow modeling by Fiering (1964). An interesting example of the reversal of space and time coordinates in conventional PCA has been given by Brier (1968).

All of these applications are basically concerned with quantitative descriptions of space-time variability. Such variability is represented in PCA by a partition of the variation of the element into a series of orthogonal spatial functions and corresponding amplitudes which vary only in time.

According to the authors of the references cited, there are two major advantages of PCA over such other possible expansions. The first is that under given assumptions principal components are the most efficient orthogonal

functions that can be chosen—that is, it can be shown that no other set of orthogonal functions can be chosen which explain a greater portion of the combined space-time variance than the empirical orthogonal functions or principal components. Second, since they are “natural” in the sense that no analytic functions are involved, the possibility that they may be interpreted physically is suggested (Sellers 1968, Stidd 1967). For purposes of this article, it must be pointed out that there are two assumptions inherent in the mathematical demonstration of the maximum efficiency of principal component analysis. The most basic assumption is that the statistical behavior of the physical quantity is stationary. This assumption is, of course, common in statistical theory, but because of changes in the behavior of the atmosphere over the globe and on the diurnal and annual time scales, it is probably not satisfied.

A further assumption is that of the homoscedasticity of the error term. That is, the mean-square error is distributed over the entire sample (space and time) in the fashion of classical least squares; and therefore the error variance does not depend, in particular, upon the magnitude of the element represented. (The statistical term *scedastic* is synonymous with variance, and the term *homoscedastic* therefore means a constancy of variance—in common usage in regression theory it denotes constant error variance.)

The relative variability of the element to be represented by a given spatial function is of fixed pattern but arbitrary sign, or sense, depending upon the sign of its amplitude. This “flip-flop” feature is a result of the fact that the spatial functions are the eigenvectors of the correlation or covariance matrix, and are therefore symmetric functions in the same sense that correlation function is

*The National Center for Atmospheric Research is sponsored by the National Science Foundation.

symmetric. Although recognized by some workers as a disadvantage (Mitchell 1960), this flip-flop feature is seen by Sellers as an advantage—"... it allows the signs of the anomalies (of monthly total precipitation over an area) to go either way." However, it should be recognized that the atmosphere may not behave in such a homoscedastic fashion and that anomalies of one sign for example (dry) may not behave in similar fashion to those of the other sign (wet).

Indeed, that such might be the case is suggested by a study of Hoyt and Langbein (1944) on variations in streamflow. Based on a study of 32 yr of annual (water-year) runoff from streams distributed over the 48 States, they point out that "... the extent of unbroken areas of deficient streamflow (below the 25 percentile) seems generally greater than that of unbroken areas of excessive streamflow (above the 75 percentile)." Taking the 5 wettest years, roughly 60 percent of 48 States had runoff in the upper quartile, and in the 5 driest years, 75 percent of the region was in the driest quartile. The application of PCA to the joint space-time variations in streamflow would in this instance not be optimum in the sense that the principal component (spatial) functions would be, morphologically, a compromise between wet and dry patterns.

The purpose of this study is to investigate the heteroscedastic behavior of some meteorological fields—specifically precipitation and streamflow—and to suggest nonparametric or distribution-free methodology capable of detecting such behavior.

2. METHODS

The most commonly used statistical techniques that are nonparametric or distribution-free are the rank-order methods. In general, these do not, except for the application of the rank correlation coefficient by McDonald and Green (1960), seem to have been applied in meteorological situations. In time series analysis, extensions of these techniques, known as runs tests, have received much more attention.

Rank-order methods, although possessing the advantage of not depending upon the underlying probability distribution functions of the variables, are frequently criticized as being less powerful statistically than the so-called parametric methods. However, it should be pointed out that the considerations of statistical power are always based upon classical alternatives—testing the null hypothesis of randomness against alternate hypotheses of linear trend, first-order Markov correlation function, etc., as examples. Against unconventional alternatives, rank-order methods may prove to be more powerful than the parametric methods. The work reported on below will attempt to demonstrate this through a particular application of rank-order techniques.

The following paragraphs outline the use of a rank-order table useful in analyzing joint space-time depend-

	space (station)					SUM
	1	2	3	...	J	
1	R_{11}	R_{12}	R_{13}	...	R_{1J}	$\Sigma R_{1.}$
2	R_{21}	R_{22}	R_{23}	...	R_{2J}	$\Sigma R_{2.}$
.
time (year)
.
.
I	R_{I1}	R_{I2}	R_{I3}	...	R_{IJ}	$\Sigma R_{I.}$
$\Sigma R_{.1} = \Sigma R_{.2} = \Sigma R_{.3} = \text{etc.}$						$\Sigma R_{..}$
$= \frac{I(I+1)}{2}$						$\Sigma \bar{R}_{..} = \frac{1}{2} J(I+1)$

$$\text{Total sum of squares} = J \cdot \frac{I(I+1)(2I+1)}{6}.$$

A dot in place of a subscript indicates summation over that subscript.

FIGURE 1.—Layout of rank table used in the analysis of space-time variability of streamflow and precipitation data.

ence. Consider figure 1. Data for each station, 1, 2, ..., J for observing times 1, 2, ..., I , are ranked according to magnitude so that each column contains the first I integers. (It makes no difference, of course, if the largest or smallest value is ranked 1, etc., as long as all columns are done the same.) Then for each time, 1 to I , the ranks appearing are summed, that is, row-wise totals are obtained. The rank-order table thus constituted has the following characteristics. First, since the first I integers appear J times, the total sum of squares of the entries in the table is simply an exact function of I and J . Furthermore, since the column-wise sum of squares must be the same for all columns (all columns contain the first I integers), the total sum of squares is partitioned into within row sum of squares (WRSS) and between row sum of squares (BRSS). If all or some of the ranks accorded a particular observation time tend to be the same, then the WRSS value will be less than the expected value for random permutation of the J, I integers. This fact leads to the definition of the coefficient of concordance (Kendall 1962)

$$W \equiv \frac{\text{BRSS}}{\text{TSS}} = \frac{J \sum_I (\bar{R}_{I.} - \bar{R}_{..})^2}{\sum_I \sum_J (R_{IJ} - \bar{R}_{..})^2} \quad (1)$$

in which a dot indicates summation over the subscript

it replaces. This quantity is analogous to a multiple correlation coefficient in parametric statistics. If the J stations are all in agreement on the ranking of the observations, the expected value of the coefficient of concordance is unity; if the ranks are permuted randomly then $E[W]=0$. The sampling theory for W on the null hypothesis of random permutation of the J rankings has been worked out. (See Edwards 1967 for a full discussion of the rank table technique and related matters.) The quantity

$$F' = \frac{(J-1)W}{1-W} \quad (2)$$

is distributed as the F -variance-ratio with $I-1$ degrees of freedom for the numerator and $(J-1)(I-1)$ for the denominator.

Second, the ranking technique allows comparison of station data regardless of what that data's probability distribution function happens to be. In the analysis of variance performed on the rank table there is no need, therefore, that the column (station) data come from the same probability distribution function, as is necessary in conventional analysis of variance.

Third, the row-wise sums, subsequently referred to as rank sums, possess features which are convenient for testing the null hypothesis of random rank permutation. For I and J given, all rank sums must be $J \leq R_i \leq IJ$. Application of the central limit theorem allows the following statement about the distribution of the rank sums. Since, in the null hypothesis of random rank permutation, each J set of I integers is identically and independently distributed, the rank sums R_i in the limit for I and J large will be identically (but not independently) and normally (Gaussian) distributed with the expected value $\frac{1}{2}J(I+1)$ and variance $J \frac{(I+1)(I-1)}{12}$. By ranking the

rank sums and plotting the R_i on Gaussian probability paper a check on the observed distribution of the rank sums is available. The coefficient of concordance, (1) above, will be proportional to the slope of the curve joining the points on such a diagram; however, such a procedure, as will be clear in the next section, allows in addition an assessment of the homoscedasticity of the coefficient.

The preceding analysis gives information on the degree and characteristics of the spatial coherence of the element observed at the J stations. In order to test nonrandom behavior of the ranks in time, a number of possibilities are present. The one chosen for this study was to use the ranks of the rank sums as the basic time series. The test, therefore, considers the temporal behavior of the combined J stations. The particular test is a version of the runs tests (Kendall and Stuart 1966, chapter 45) used to test randomness of ranked data. However, in an effort to circumvent some of the unsatisfactory aspects of runs tests, the analysis here uses some theory on random permutations by Gray (1967). The basic problem is:

what are the probabilities of observing sequences of runs of A things in a total of A and B things if the A and B things are arranged randomly? A run of length k is defined as k consecutive occurrences of the A things regardless of where such a sequence occurs. The theory by Gray permits the computation of the exact probability of getting any of the possible combinations in a series of random permutations of A things in $A+B$ items. Let a_k be the number of runs of the A things of length k so that

$$\sum_{k=1}^A k a_k = A,$$

and let the total number of runs

$$\sum_k a_k = a.$$

Then the probability of getting exactly run arrangement L is

$$P(L) = \frac{(B+1)!}{\left\{ \prod_{k=1}^A a_k! \right\} \{B-a+1\}!} \cdot \frac{A!B!}{(A+B)!} \quad (3)$$

For example, consider 5(= A) things in a total of 25(= $A+B$) things. One possible arrangement is for all 5 items to occur in isolation: another is for one run of 2 (2 of the 5 things occurring together) and 3 in isolation, etc. In notation to be used, the following sequences are all of those possible considering $A=5$ things: (# of runs of 1/, # of runs of 2/, # of runs of 3/, etc.) so (5/0/0/0/0), (3/1/0/0/0), (1/2/0/0/0), (2/0/1/0/0), (0/1/1/0/0), (1/0/0/1/0), and (0/0/0/0/1) are the seven possible combinations. If we now consider that the probabilities of each run pattern allow a stratification of the patterns by probability, from highest to lowest, then a method whereby a level of probability can be attached to any observed pattern L' is suggested. The quantity

$$1 - \frac{P(L')}{\sum_{sup P(L)} P(L)}$$

then gives the observed probability of getting a sequence more nonrandom (of a lower probability) than the one observed on the null hypothesis of random permutation of the (A and B) things. Clearly, the probability of observing the last run pattern above in a series of random events is much less than observing the first. For instance, the results of the example are: $P(3/1/0/0/0)=0.450$; $P(5/0/0/0/0)=0.383$; $P(1/2/0/0/0)=0.075$; $P(2/0/1/0/0)=0.075$; $P(0/1/1/0/0)=0.008$; $P(1/0/0/1/0)=0.008$; and $P(0/0/0/0/1)=0.000+$. There is certainly a relationship between the calculated probability and the degree of subjective nonrandomness in the seven sequences—those exhibiting a persistence or “clumping” of the A things have the lowest probability. Furthermore, if we had observed, say, a $P(0/1/1/0/0)$ sequence in some trial,

TABLE 1.—Summary of data nets. (See appendix for details.)

Data net	No. of years	No. of stations	Source	Averaging period
Larger basins Streamflow	1908-1966 (59)	12	USGS Water Supply Papers	Water-year
Smaller basins Streamflow	1908-1966 (59)	12	USGS Water Supply Papers	Water-year
Net I Precipitation	1908-1966 (59)	20	USWB Bulletin W	Various seasonal & Nov.-May totals
Net II Precipitation	1893-1966 (73-74)	18	USWB Bulletin W	Various seasonal & Nov.-May totals

then the significance level (hypothesis of random sequence) would be $1 - (0.450 + 0.383 + 0.075 + 0.075) = 0.017$.

3. DATA

Four sets of data were analyzed by the methods outlined in the previous section. Table 1 presents relevant information for the two streamflow and two precipitation nets chosen. There were a few cogent reasons why these particular hydrometeorological variables were chosen. Streamflow is one indicator of climate that is of very practical importance in addition to its intrinsic value as a climatic variable. Stochastic modeling of streamflow is a vigorous field of research currently, and a great deal of effort has gone into effective methods of analysis of hydrologic data (Fiering 1967, Subcommittee on Hydrology 1967). Precipitation is another related variable that has received a good deal of attention from statisticians.

Moreover, an important feature of precipitation data is the statistical fact that the ratio between time-averaged amounts for two stations is a conservative quantity dependent only, apparently, upon storm type (Williams and Peck 1962). This feature is used to assess the time stationarity of station records and to adjust nonstationary records when feasible. The technique, commonly referred to as double-mass analysis, is described in many hydrology texts, for example, Linsley et al. (1949); the particular details in this investigation are relegated to the appendix. Another important climatic variable, temperature, does not follow this behavior of conservative ratios. As a consequence, adjustment to achieve stationarity of temperature records is a difficult if not impossible task (Mitchell 1961).

Maps indicating the location of the stream runoff basins and the rain-gage stations used in the two precipitation nets are shown in figures 2 and 3. The particular stations shown are the result of a necessary compromise between the demands for a regular spatial distribution, reasonably long record length, and stationary record. Two sets of both streamflow and precipitation were used to obtain some idea as to sampling variations due to arbitrary station location. The streamflow data is totaled by water year, October 1 to September 30, and the precipitation data totaled in various ways as indicated—generally by season.

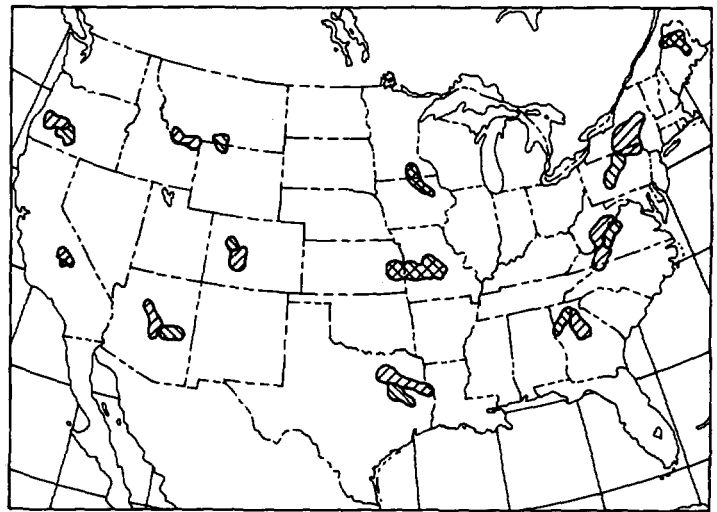


FIGURE 2.—Map showing the location of the basins used in the analysis of streamflow data. Cross-hatching differentiates the two nets used.

4. RESULTS—SPATIAL VARIATION

Figures 4 through 6 were constructed by calculating the plotting position (probability) from $\text{prob}(m) = m/(n+1)$, where m is the rank of the rank sum R_i for each of the n years. The quantities $\text{prob}(m)$ and R_i were then plotted on Gaussian probability paper. Thus, for the 12 smaller streamflow basins (fig. 4), the rank sums ranged between 628 for the driest and 190 for the wettest of the 59 yr, 1908-1966. Recalling the theory of the distribution of rank sums, for the null hypothesis that the data are characterized by $59(=I)$ samples of $12(=J)$ independent series, the mean rank sum would equal $1/2(12)(60) = 360$ with standard deviation $(60 \cdot 58)^{1/2} = 59$.

12

However, instead of the sampling variations of these quantities, the sampling theory of W can be used to assess the significance of the average slope of the points shown. Table 2 gives the appropriate results for $F'(W)$ for all data sets considered. For the smaller basins, for which figure 4 is appropriate, we see that the coefficient of concordance is far in excess of significance of the 5 percent level, permitting the conclusion of significant spatial nonrandomness of the stream runoff data. In table 2, we tentatively assume that all years are independent—an assumption that will be checked and modified if necessary in the next section.

More important, however, is the indication of a change of slope (or skewness) of the rank sums between the dry years (left-hand side or $\text{prob}(m) < 0.50$) and the wet years (right-hand side or $\text{prob}(m) > 0.50$). This slope change, if significant, is quantitative support of Hoyt and Langbein's (1944) analysis and is indicative of a heteroscedasticity in the spatial coherence of the streamflow data. Since the dry years exhibit a greater slope than the wet years, the

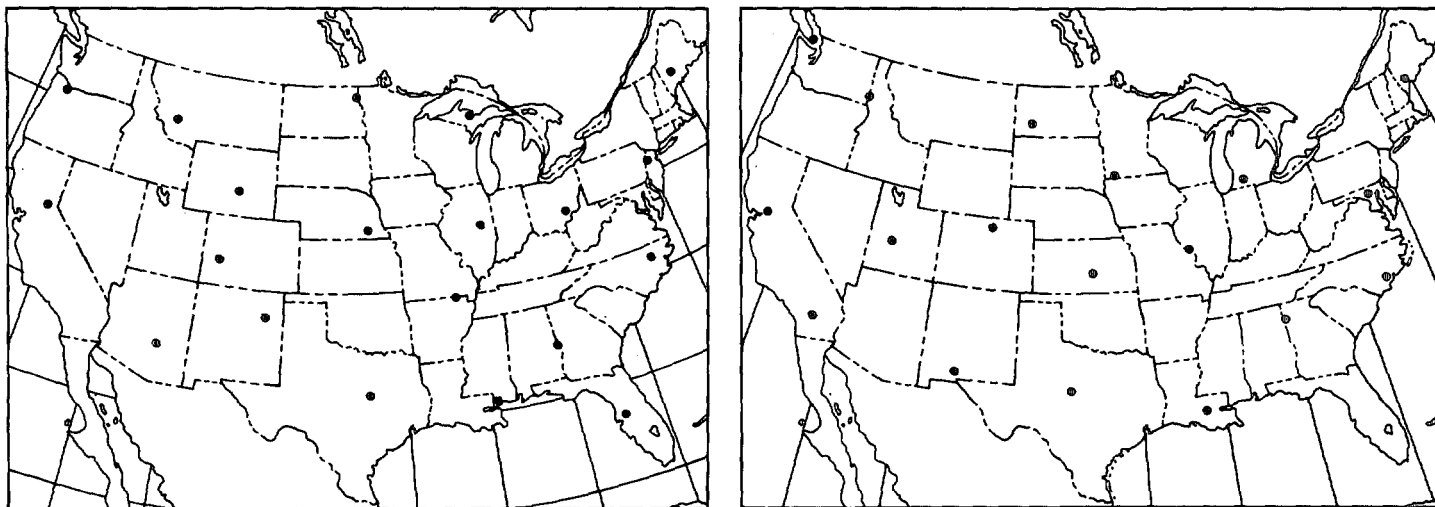


FIGURE 3.—Map showing the location of the stations used in the analysis of precipitation data. Net I, 20 stations, left; and net II, 18 stations, right.

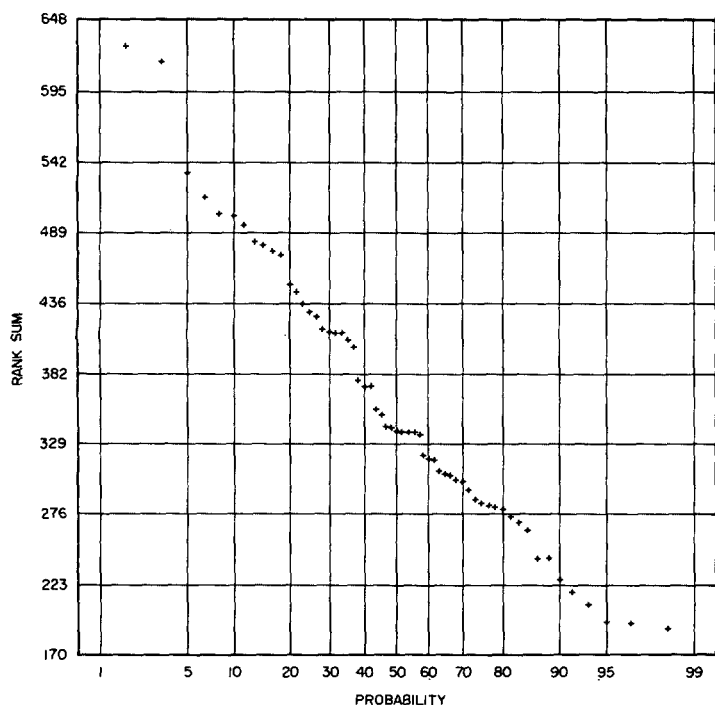


FIGURE 4.—Distribution of the rank sums for the network of smaller streamflow basins plotted on Gaussian probability paper.

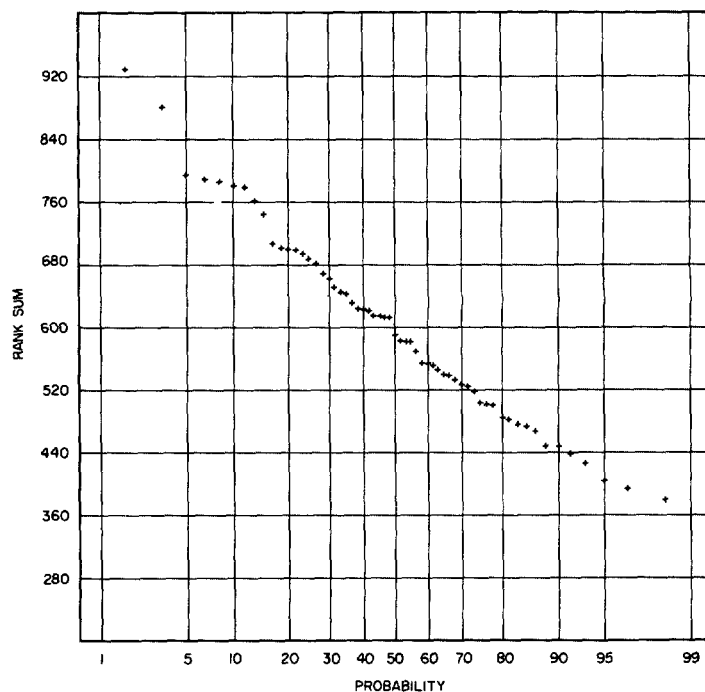


FIGURE 5.—Same as figure 4, but for November through May precipitation totals for net I.

analysis implies that the former exhibits a greater spatial coherence than the latter. Physically, this result can be interpreted that "drought" has a greater and/or stronger horizontal (spatial) correlation than wet years or "non-drought," as Hoyt and Langbein suggest. An analogy in more familiar terms would be a correlation scattergram for two variables x and y in which the scatter is markedly different in one part of the diagram from another; or, in other words is a function of x or y .

The distribution of rank sums for the 12 larger basins is not shown because it is similar to figure 4. Figures 5 and 6 give the distribution of rank sums for the two precipitation nets for November through May totals. The same behavior is noted, particularly for net II, which covers a longer time period than net I. Thus, the tendency for dry years to be spatially more coherent than wet years is a feature of winter half-year precipitation as well as of streamflow.

The foregoing analysis was performed on November-May precipitation totals for two reasons: first, because

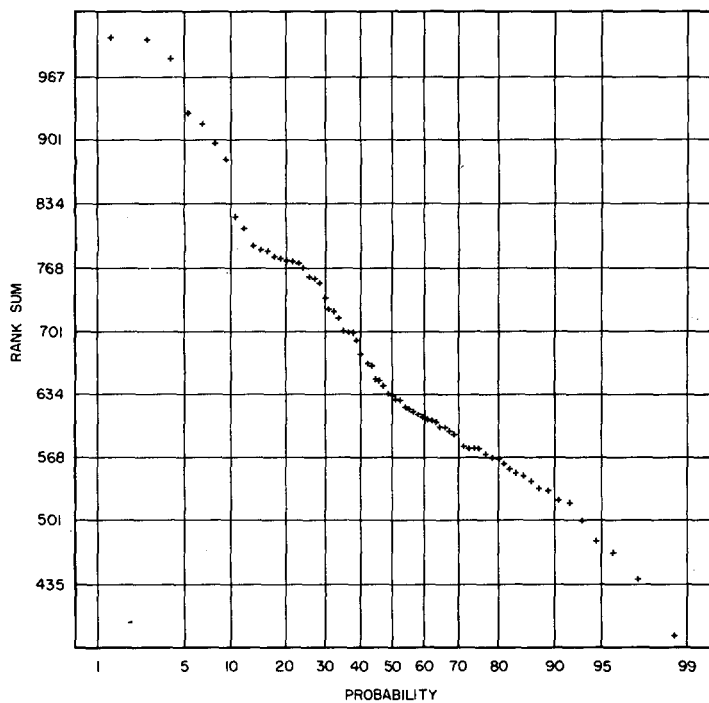


FIGURE 6.—Same as figure 4, but for November through May precipitation totals for net II.

TABLE 2.—Parameters of the distribution of rank sums

Data set	F'	γ_1	Cornu ratio
Net I Nov.-May	2.76* (1.40)	0.400 (0.493)	0.807 (0.755)
Net II Nov.-May	2.33* (1.36)	.623* (.458)	.803 (.758)
Net II Sept.-Nov.	2.47* (1.36)	.496* (.458)	.802 (.758)
Net II Dec.-Feb.	1.70* (1.36)	-.245 (.458)	.732* (.758)
Net II Mar.-May	2.31* (1.36)	.343 (.458)	.808 (.758)
Net II June-Aug.	1.50* (1.36)	-.097 (.458)	.779 (.758)
Streamflow (larger)	3.60* (1.41)	.480 (.493)	.800 (.755)
Streamflow (smaller)	3.57* (1.41)	.486 (.493)	.815 (.755)

Significance levels ($P=0.05$) on the null hypothesis of random rank permutation shown in parentheses—values with asterisks are those for which the null hypothesis is rejected

this roughly half-year time period should be relatively free of precipitation from small-scale summer convective storms and hurricanes, and most representative of that from extratropical cyclones. Further, this time interval produces the highest correlation of the precipitation and streamflow rank sums (rank correlation=0.69 for net I November–May precipitation and the 12 smaller basins). It is, of course, of interest to investigate the behavior of the rank sums and W when shorter time intervals are used to total the precipitation. Figure 7 shows the variation in F' , equation (2), by season (conventional 3-mo seasons) for nets I and II. Also shown are the values of F' ($P=0.01$) on the null hypothesis of random rank permutation. The summer convective season exhibits the lowest values of $F'(W)$, which are barely significant at the 1 percent level. It is perhaps surprising to find that summer rainfall, which at the majority of the stations must be overwhelmingly

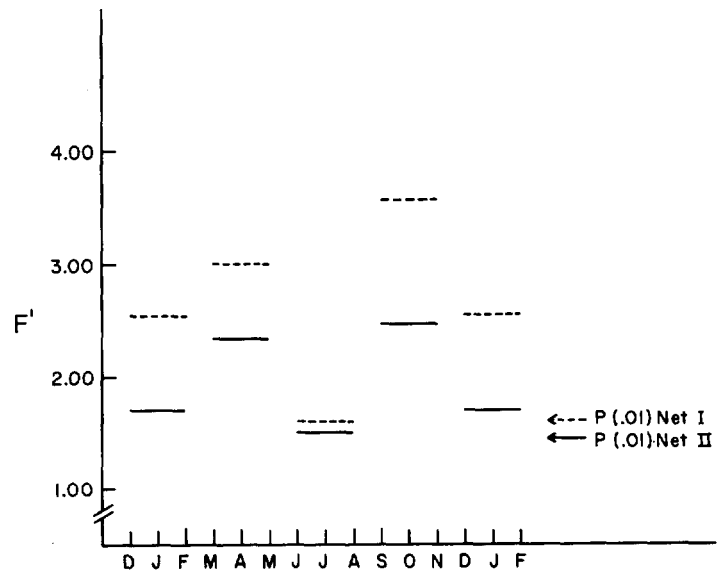


FIGURE 7.—Seasonal variation in the spatial coherence of precipitation as measured by the observed F' -variance ratio, equation (2). The 1 percent significance levels on the null hypothesis of no spatial coherence are shown.

influenced by convective showers, still possesses sufficient spatial coherence to produce significant values of the coefficient of concordance. Highest F' values are for the transition seasons. The meteorological interpretation of this result seems obscure, but one factor that comes immediately to mind is the problem of rain-gage catch for snow *versus* liquid precipitation. Snow catch is apparently much more affected by wind than is liquid precipitation catch (Bruce and Clark 1966), and a majority of the stations may be exhibiting a measurement error owing to this factor.

The problem of the significance of the variation in slope, or the heteroscedasticity of the rank sums, can be attacked in a relatively straightforward manner. For random rank permutation, the rank sums for large I and J approach a Gaussian distribution, as pointed out in section 2. For a sample of N rank sums drawn from a Gaussian distribution with sample variance \hat{m}_2 and third moment \hat{m}_3 , the standardized skew coefficient has an expected value of 0 and a standard error of $\sqrt{6/N}$. The t -distribution on the standardized skew coefficient $\gamma_1 = \hat{m}_3/(\hat{m}_2)^{3/2}$ may therefore be used.

Table 2 presents the γ_1 skewness statistic for the various data sets with the appropriate one-tailed values of γ_1 at the 5 percent significance levels (Pearson and Hartley 1962, table 34B): the one-tailed test was deemed appropriate because of the *a priori* expectation of positive skewness suggested by Hoyt and Langbein's (1944) work. Also in the table are the Cornu ratios, the ratio of the mean deviation to the standard deviation, for the rank sums of all the data sets (O'Brien and Griffiths 1967). This statistic is a measure of kurtosis, or the fourth moment of the distribution. For the Gaussian distribution, the

expected value of the Cornu ratio is 0.800; values of the critical Cornu ratio, again from a one-tailed test on the 5 percent significance level, are in parentheses (Pearson and Hartley 1962, table 34A). A one-tailed test was chosen here because the finite limits of J and IJ for the rank sums suggests the observed distribution might possess a negative coefficient of kurtosis.

The skew coefficient values, γ_1 , for the November–May and September–November totals of the net II precipitation data exceed the 5 percent limits. Quite noticeable is a positive correlation between positive skewness and the value of F' itself. Thus, the greater the overall spatial coherence the greater the skewness in the distribution of the rank sums. This fact and visual examination of figures 4 to 6 suggests that it is the dry years that are producing the significant values of W and simultaneous skewness in the rank sums. That is to say, the right-hand (or wet) half of the rank sum distribution curves has a slope which is characteristic of near-random permutation of the ranks, whereas the left-hand (or dry) side produces a larger W and indicates a coherence in those rankings. A study to verify this conclusion was made by stratifying the 73 yr of November–May net II precipitation into two separate samples—wet and dry—based on the rank sums shown in figure 6. For each sample, the rank table computations and coefficient of concordance were computed separately. The W value for the wet years was 0.24, not significant at the 5 percent level; and for the dry years, 0.85, which is significant at that level. It appears, then, that a stronger spatial coherence of the dry years is the basic phenomenon detected in figures 4–6.

Another interesting aspect of table 1 is the apparent relationship between the skew coefficient γ_1 and the Cornu ratio. The two sets of precipitation data, winter and summer, having small negative skew coefficients also have anomalously low values of the Cornu ratio (the winter precipitation Cornu ratio differs significantly from the Gaussian distribution value of 0.800 at the 5 percent level). This fact also tends to confirm the conclusion that the stronger the spatial coherence the more non-Gaussian is the distribution of the rank sums.

One further investigation into the differences between the wet and dry years was made. The separate rank correlation matrices for the dry- and wet-year samples were compared by calculating the quantity

$$\Delta(J) = \frac{1}{17} \sum_{L=1}^{18} [\hat{\rho}(J, L)_{dry} - \hat{\rho}(J, L)_{wet}], \quad L \neq J, \quad (4)$$

or, in other words, the columnar average of the differences, neglecting the diagonal ones, between the correlation matrix for the dry years and that for the wet. The $\Delta(J)$'s thus are equal to the average difference between the correlations of a station J and all other stations for the dry years over that for the wet years. Figure 8 shows the $\Delta(J)$ quantity plotted for the November–May net II precipitation data. It is quite evident that the stations in

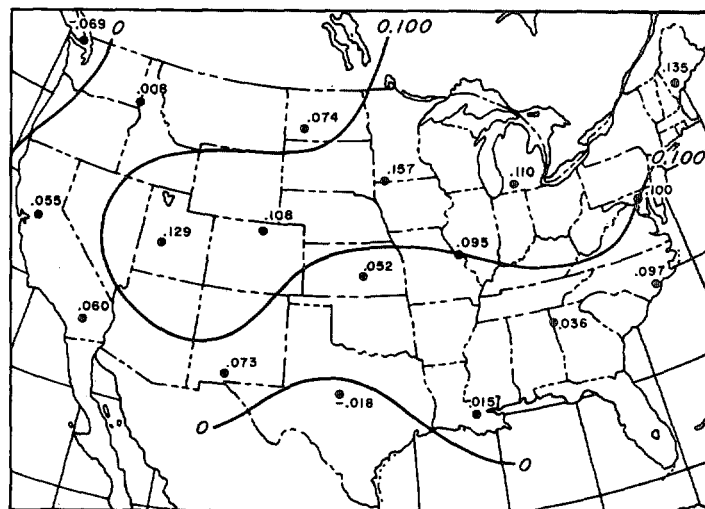


FIGURE 8.—Map of the average difference in the spatial correlation coefficients of the dry minus the wet years, equation (4).

the interior of the country and along the east coast are better intercorrelated in dry than in wet years, while stations along the Pacific and Gulf Coasts show little or no difference. This pattern may have physical significance because the Pacific Ocean and Gulf of Mexico are the two major source regions for moisture in the continental United States. The difference in correlation might be thought of as a kind of continentality effect. The dry years, according to this interpretation, do not exhibit a higher coherence on a larger horizontal scale, but rather are simply more homogeneous or more “in agreement” in the interior of the continent.

5. RESULTS—TIME VARIATION

The results of the previous section suggest that in analogous fashion to the different spatial behavior of wet and dry years there may also be different behavior in the time dimension. The theory on probability of runs outlined in section 2 was applied to the four data nets to examine this possibility. An arbitrary *a priori* selection was made by choosing to compare the behavior of the driest and wettest quartiles. Thus, for example, for the 59-yr streamflow and precipitation samples, the 14 driest and the 14 wettest years as determined by the rank of the rank sums were used. From the theory, equation (3), the probabilities of all possible combinations of 14 things distributed randomly in 59 were computed (for net II, 18 things in 73). The run combination with highest probability was the (10/2/0) combination with $P=0.193$; the second highest was (8/3/0), $P=0.166$; and the third (9/1/1/0), $P=0.110$. The cumulative probability of observing a run sequence with a lower probability than these three is thus $(1.0 - (0.193 + 0.166 + 0.110)) = 0.531$, approximately the 50 percent level.

TABLE 3.—Results of runs test on ranks of rank sums

	Dry quartile	P ()	Wet quartile	P ()
Large streamflow net.....	(5/3/1/0)	0.22	(7/2/1/0)	0.42
Small streamflow net.....	(4/2/2/0)	.05	(5/3/1/0)	.22
Net I Nov.–May precip.....	(4/5/0)	.06	(8/3/0)	.64
Net II Nov.–May precip.....	(11/2/1/0)	.75	(10/4/0)	.64
Net II Sept.–Nov. precip.....	(10/2/0/1/0 . . .)	.19	(8/2/2/0)	.25
Net II Dec.–Feb. precip.....	(10/4/0)	.64	(10/2/0/1/0 . . .)	.19
Net II Mar.–May precip.....	(10/4/0)	.64	(8/2/2/0)	.25
Net II June–Aug. precip.....	(11/2/1/0)	.75	(9/1/1/1)	.14

Table 3 presents a summary of the results of this runs test. For the four data nets and for the dry and wet quartiles, the particular observed run sequence is given together with the probability, on the null hypothesis of randomness, of observing a less probable sequence. Only for one of the eight combinations should the null hypothesis of randomness be rejected at the 5 percent level (the 12 smaller streamflow basins, dry years), although one reaches the 6 percent level (November–May total precipitation net I, dry years). However, the probability levels for the various samples seem quite unstable. Moreover, the relative behavior of the dry and wet quartiles is not very consistent.

It should be pointed out that water-year streamflow amounts, for some individual basins at least, are known to be nonrandomly distributed in time (Yevjevich 1964). We might expect, therefore, that the probability levels for streamflow would be relatively low (that is, greater deviation from randomness), and this seems borne out by the results. Of more interest here are the *differences* between the dry and wet years for streamflow. For both streamflow nets, the probability levels are lower for the dry than for the wet years, suggesting a heteroscedasticity in time. Further investigation of this possibility seems warranted, particularly because of the implications such an effect would have for stochastic streamflow modeling.

The temporal behavior of streamflow from individual basins is known to vary widely depending upon various hydrological factors (Langbein et al. 1949). It is perhaps unrealistic to analyze the joint behavior of different basins spread over such a heterogeneous region as the 48 conterminous States. However, water-year streamflow is a climatic indicator related to precipitation and one that possesses variability on many spatial and temporal scales. Obviously, streamflow is correlated spatially; and, therefore, regardless of the hydrological details, the joint space-time behavior in a gross sense is worthy of study. Because precipitation and streamflow are related, one other aspect of their temporal behavior on a continental scale is given by the data in table 4.

In this table, a summary is given for 14 wettest and driest years in the 1908–1966 period determined from the rank sums of the smaller streamflow (SF) net and net I November–May precipitation. Eight years (1925, 1926,

TABLE 4.—Comparison of the 14 wettest and driest years from various data nets

Twelve smallest basins Net I Nov.–May precipitation 1908–1966				Twenty basins (Langbein 1968) Net II Nov.–May precipitation 1896–1966			
14 wettest years		14 driest years		14 wettest years		14 driest years	
SF	P	SF	P	SF	P	SF	P
1952	1949	1934	1934	1909	1903	1934	1934
1908	1922	1931	1925	1903	1949	1931	1931
1916	1952	1959	1963	1916	1922	1941	1925
1909	1908	1939	1911	1907	1935	1959	1911
1912	1938	1930	1931	1908	1897	1940	1955
1943	1937	1940	1940	1897	1948	1959	1918
1920	1945	1955	1918	1902	1958	1925	1963
1927	1929	1954	1913	1906	1919	1930	1964
1945	1953	1964	1955	1965	1912	1924	1959
1951	1958	1925	1930	1952	1923	1933	1902
1922	1965	1956	1917	1922	1900	1955	1959
1958	1912	1966	1926	1927	1944	1937	1956
1921	1919	1941	1956	1901	1929	1926	1926
1928	1920	1926	1910	1912	1953	1911	1966

TABLE 5.—Theoretical and observed frequencies of the 14 wettest and driest years before and after 1922

	1908–1966 (table 4)		1896–1966 (table 4)	
	SF	P	SF	P
Observed frequency—dry years before 1923...	0.000	0.357	0.071	0.214
Theoretical frequency, random distribution...	.540	.540	.380	.380
Observed frequency—wet years after 1922....	.500	.642	.214	.571
Theoretical frequency, random distribution...	.746	.746	.620	.620

1930, 1931, 1934, 1940, 1955, and 1956) appear in both lists for the dry years, and 7 yr (1908, 1912, 1920, 1922, 1945, 1952, and 1958) for the wet years.

Also in table 4 are the 14 wettest and driest years in streamflow, 1896–1966, taken from Langbein (1968), together with the comparable years for the net II November–May precipitation (P). Here, 8 yr occur in both “dry lists” and only 4 yr in the “wet lists.” This disparity, which is apparently reinforced by considering the larger sample, is a manifestation of a possible heteroscedasticity in the wet and dry years—the correlation coefficient between streamflow and precipitation (here=0.69) is produced largely by the correspondence in the dry years.

An important consideration arises when a count is made of the 14 driest years occurring before 1923 and the number of wet years occurring after 1922. A comparison is given in table 5 of the actual frequency and the theoretical frequency that would be expected if 14 yr occurred at random throughout the 59 (or 71) years. The extreme streamflow years can be seen to be highly biased—nearly all of the 14 wettest years in each group occurred before 1923 and the driest after 1922. Such a bias is only sug-

gested, however, in the precipitation figures. No test of significance is employed here, mainly because the 1922–23 division is an arbitrary one and was picked after an examination of the data.

This result, so evident to the subjective eye when examining streamflow data for most of the basins in the United States, must be a result of two hydrologic facts. Because, to a first approximation, streamflow is a residual of two larger compensating quantities—precipitation and evapotranspiration—it may exhibit a different time dependence than either of the larger quantities. Therefore, if we assume November–May precipitation to be randomly distributed in time and water-year streamflow is observed to be nonrandom, exhibiting persistence of some other longer term time dependence, then obviously the evapotranspirative process could be instrumental in producing the latter. The small amount of stored carry-over water contained in water-year streamflow amounts cannot account for *all* of the observed temporal streamflow dependence (see Yevdjovich 1964 for a discussion and data).

The other certain and important contributor to the observed nonrandomness is the nonstationary component in streamflow because of manmade changes in the natural stream basin, increased consumptive use, ground water depletion, etc. The critical question, how large and how important this component is, is not answerable; but most hydrologists would agree that it would account for most of the observed differences in the streamflow and precipitation groupings in figure 5. The relevant consideration for this investigation, however, is the degree to which the presence of this nonstationary component in streamflow influences the heteroscedasticity of the correlation between precipitation and streamflow. To be specific, if natural and stationary streamflow data were available, would the ranking results as in table 4 indicate a greater or lesser disparity in correspondence in dry and wet years? Because of the importance of this consideration, an attempt should be made to test for this time-wise heteroscedasticity—perhaps initially by using local data and ignoring the spatial problem.

6. CONCLUSIONS

The rank-order statistic methodology that has been suggested can be used to investigate joint space-time dependence of any scalar quantity on any scale. The advantages of the scheme seem to be 1) it is nonparametric and permits employment of data from different and unknown probability density distributions, and 2) it allows the assessment of the degree of homoscedasticity in the spatial and temporal associations in the data.

The disadvantages are 1) the scheme does not use analytic functions of any kind and furthermore it does not suggest in any obvious fashion how the technique of principal components can be modified to account for any

heteroscedasticity present, and 2) the coefficient of concordance and the distribution of rank sums are essentially a measure of agreement. The converse, disagreement, is not well defined by the analysis, and as a consequence, the tests of W on the null hypothesis of random permutation of the ranks cannot be very powerful against certain alternatives. For example, if two sizeable fractions of the J stations each have ranks which tend to be similar within each fraction but which average very high in one fraction and low in the other, or *vice versa*, the within row sum of squares would be indistinguishable from the maximum value present when the ranks for each station are permuted randomly.

The temporal analysis used is also insensitive to various nonrandom alternatives, for example, cyclic variations.

The joint space-time variation of precipitation over the 48 States has been shown to possess a basic heteroscedasticity. Periods with deficient precipitation tend to be more coherent spatially than periods with excessive precipitation. The same effect is seen in stream runoff variability. Further, the type of heteroscedasticity exhibited seems to be characterized by a significant coherence in the wet periods. The meteorological interpretation of this behavior of precipitation and stream runoff must involve consideration of spatial differences in the time-averaged behavior of precipitation patterns on one hand, as contrasted with the relative precipitation-free areas on the other.

Not unlike studies of the temporal behavior of seasonal precipitation by station, no significant nonrandomness was detected in the distribution of wet and dry periods. Stream runoff, however, over the United States exhibits a distinct tendency toward a temporal nonrandomness with a suggestion that dry years are more persistent than wet ones. Unfortunately, the extent to which this persistence is caused by the nonstationarity of the streamflow data cannot be determined.

Although the analysis has achieved its purpose in pointing out limitations to the analytic-statistical representation of joint space-time variations by the use of principal components, it has not, as noted above, suggested any modification to the latter to take into account the observed heteroscedasticity. Further work is anticipated—the specification of a scedastic function, that is, distributing the error over space in the estimation of the covariance or correlation matrix is one possibility (Glejser 1969).

Finally, a quote from Freiburger and Grenander's (1965) paper, "On the Formulation of Statistical Meteorology," seems particularly appropriate:

"The statistical techniques we have discussed so far (PCA, regression theory, and discriminate functions) are all based on stochastic models of such simple structure that the underlying assumptions are often not specified explicitly. Many authors have realized that a more sophisticated description is needed to achieve realism and to provide a better understanding of the phenomena."

APPENDIX — SUPPLEMENTARY INFORMATION ON DATA AND DATA NETS

STREAMFLOW

All streamflow data were taken from the "Water Supply Papers" of the U.S. Geological Survey (1908–1966). These data are actual gaged streamflow for the water year, October 1–September 30, adjusted for any diversions and artificial regulation for which records are available. The data therefore approximate the natural behavior of the streams, which is variously termed runoff (U.S. Geological Survey 1960) or full natural flow (Linsley et al. 1949). In this paper, the term streamflow is used as a synonym for runoff—the latter is not sufficiently informative by itself and "stream runoff" seems unnecessarily awkward. The streamflow data were not subjected to record homogeneity (stationarity) tests because of the difficulty in taking into account the influence of manmade depletion on the observed records. As noted in section 5, the resulting nonstationarity of these data introduce difficulties into the analysis and extenuates any conclusions drawn therefrom.

The two streamflow nets consist of basins which, but for three, are in paired juxtaposition, or nearly so. It was necessary to include three isolated basins in both nets to achieve reasonable spatial coverage. The basins used, together with relevant information, are included in table 6.

PRECIPITATION

The task of choosing the stations to be included in the precipitation nets was an extremely time-consuming and frustrating one. The demands of record homogeneity (stationarity), location, and record length resulted in the examination of over 150 stations, about half of which were ultimately discarded. The double-mass technique was used to check record homogeneity in the following manner: groups of two or more stations in a region were chosen on the basis of their recorded station histories (U.S. Weather Bureau 1956) for intercomparison. For each group, seasonal double-mass (or cumulative total) curves were prepared for all possible combinations of the stations. The seasons were not fixed, but were adjusted to conform to local climatological seasons. For example, seasons for stations located along the Pacific coast were different than those for stations in the desert southwest, interior of the continent, etc. This procedure is more realistic than that using annual totals because of the fact that the ratio of precipitation catch is a function of storm type (Williams and Peck 1962).

In a station grouping, the probability was quite high that one or more of the stations had a recorded move or moves. Only those stations for which the double-mass curves clearly showed changes in slope (changes in ratio of precipitation catch) only at the recorded move dates were retained with one exception. When, in a group of three or more stations, a change in slope at an undocu-

TABLE 6.—Two streamflow nets and related information

Basin	Gaging station	Area (sq mi)
Umpqua	Elkton, Oreg.	3,683
Rogue	Raygold, Oreg.	2,020
Kings	Piedra, Calif.	1,694
Kaweah	Three Rivers, Calif.	418
Verde	Bartlett Dam, Ariz.	6,180
Salt	Roosevelt, Ariz.	4,981
Beaverhead	Barratts, Mont.	2,737
Yellowstone	Corwin Springs, Mont.	2,623
Colorado	Gunnison, Colo.	3,980
Roaring Fork	Glenwood Springs, Colo.	1,460
Cedar	Cedar Rapids, Iowa	6,510
Osage	Bagnell, Mo.	14,000
Sabine	Logansport, La.	4,858
Neches	Rockland, Tex.	3,539
Penobscot	West Enfield, Maine	6,600
Susquehanna	Wilkes-Barre, Pa.	9,960
West Branch Susquehanna	Williamsport, Pa.	5,682
Kanawha	Kanawha Falls, W. Va.	8,367
James	Carterville, Va.	6,240
Oconee	Dublin, Ga.	4,400
Chattahoochee	West Point, Ga.	3,550

mented date was evident which could unequivocally be attributed to just one of the stations, the group was retained. When internal consistency among the stations could not be established for the undocumented inhomogeneity, or if the double-mass curves possessed gradual curvature, the entire group of stations was dropped from further consideration. Table 7 lists those station groups for which, with reasonable certainty, one or more of the stations possess a homogeneous record, or for which one station required adjustment.

To illustrate the type of analysis involved in the assessment of station homogeneity (stationarity), figure 9 is presented. The key station here is Woodstock College, Md. This cooperative observing station is an ESSA-USWB climatological reference station, and its record has been the subject of previous investigations of climatic change, for example, Landsberg et al. (1959). For the winter season, here November through March, Woodstock's (station #96) accumulated precipitation is plotted against the comparable sum of Fallston, Md. (station #95), and Dale Enterprise, Va. (station #25). According to official station histories, none of the three stations were moved in the 1890–1953 interval, although the documentation on Woodstock may not be complete. Landsberg et al. (1959) refer to undocumented moves in March 1901 and January 1914. A straight edge laid along the plotted points makes possible the detection of changes in slope in the curve about 1914 and 1940. The latter break is inferred to be characteristic of the Dale Enterprise record because it always appears when that station is contrasted with the others, and does not appear in the plots of the other stations. Although the change in slope about the 1914 point is small, it is considered significant because of the *a priori* information concerning a station move. Appropriate corrections on the change in slope 1890–1914, 1914–1940 were therefore made in the November through March Woodstock record.

TABLE 7.—Groups of stations used in checking homogeneity of records

Net	Station	Dates used	Recorded moves	Double-mass check
Net II	Olga, Wash.	1/90-	none	OK
	Sedro Woolley, Wash.	8/96-	7/43, 7/46	
Net I	Vancouver, Wash.	10/07-	none	OK
	Headworks, Oreg.	1/10	8/52, 3/55	
Net II	*Davis (Ag. Coll.), Calif.	1/72-	6/08	OK
Net I	Colfax, Calif.	2/70 -	none	OK
	Grass Valley, Calif.	9/72-	10/08, 11/12	
	Ft. Ross, Calif.	11/74-	3/07	
	Knights Landing, Calif.	1/78-3/54	3/54	
Net II	*Indio, Calif.	1/72-	1/06, 1/19	1 break, adjusted before 1906
	Tustin-Irvine, Calif.	1/77-	1/14, 2/20, 4/41	
	Yuma, Ariz.	1/86-	none	
Net II	Moscow, Idaho	10/93-	"minor," 12/60	OK
	Orofino, Idaho	7/06-	none	
	Kellogg, Idaho	3/05-12/60	none	
Net I	E. Anaconda, Mont.	9/05-	none	OK
	Yellowstone N.P., Wyo.	9/05-	none	
Net I	Pathfinder Dam, Wyo.	1/06-	none	Not done
Net II	Fillmore, Utah	1/92-	1/38	OK
	Levan, Utah	5/89-	12/14, 8/15, 3/55	
Net II	Ft. Collins, Colo.	1/93-	12/09, 1/40, 11/56	3 breaks, adjusted to 1/10-11/39 period
	Watervale, Colo.	1/95-	none	OK
	Longmont, Colo.	11/10-	none	OK
	Hawthorne, Colo.	9/08	none	OK
Net I	*Montrose (#2), Colo.	1/06-	none (?1/11)	OK, also with 1 station sum
	Grand Junction, Colo.	1/06-	Many	OK, also with 10 station sum
	Delta, Colo.	2/07-	none (obs. change 1946)	
Net I	Pinal Ranch, Ariz.	9/95-	none	OK
	Walnut Grove, Ariz.	11/97-	11/55	
Net II	State U. (Las Cruces), N. Mex.	1/92-	"minor," 6/64	OK
	Lordsburg, N. Mex.	1/92-	4/06, 1/62	
	Elephant Butte, N. Mex.	1/08-	7/33	
Net I	Bell Ranch, N. Mex.	1/05-	none	OK, few missing months estimated from pair
	Tucumcari, N. Mex.	1/05-	3/15	
	San Jon, N. Mex.	1/07-	10/63, 8/66	
Net II	*Dickinson (Exp. Farm), N. Dak.	1/92-	none	OK
	Devils Lake, N. Dak.	1/04-	8/64	
Net I	Grafton, N. Dak.	1/92-	8/66	OK
Net II	Worthington, Minn.	1/94-	none	OK, few missing months & 11& 12/93 estimated from Bird Island & Luverne, Minn., & Sibley, Iowa
	Milan, Minn.	9/93-12/67	6/62	
	Bird Island, Minn.	1/92-12/67	3/59	
Net I	*Crete, Nebr.	8/90-	6/52, 11/61	OK
	Geneva, Nebr.	8/90-	1/19, 6/50	
Net II	McPherson, Kans.	1/90-	10/55	OK
	Hays, Kans.	1/90-	? , 01	
	Alton, Kans.	1/08-	none	
Net II	Coleman, Tex.	3/94	? , 6/10	OK, 11/93-2/94 estimated from Brownwood & Abilene
	Hewitt, Tex.	3/94-	10/65	
	Mexia, Tex.	9/04-	12/66	
Net I	Ishpeming, Mich.	1/94	8/55	OK
	Antigo, Wis.	1/94-	2/65	
Net II	Kalamazoo, Mich.	1/93-	none	OK
	Battle Creek, Mich.	1/95	4/41, 3/63	
Net I	*Urbana, Ill.	10/07-	10/49	OK
	Whitestown, Ind.	11/08-	10/60, 10/65	
	Collegeville, Ind.	10/07-	9/59	
Net II	Grafton, Ill.	1/94-	? , 1/63, 6/65	OK, 11&12/93 estimate from St. Charles, Mo.
	Cairo, Ill.	1/81-	6/42	
	Brunswick, Mo.	10/89-	none	
Net I	Pocahontas, Ark.	1/96-	none	OK
	Jonesboro, Ark.	1/96-	1/61	
	Mountain Grove, Mo.	5/01-	none	
Net II	Donaldsonville, La.	9/92-	10/65	OK
Net I	Pearlington, Miss.	9/92-	3/06, 8/46	OK
	Bay, St. Louis, Mo.	10/03-	2/39, 8/66	
Net I	Farmington, Maine	1/91-	none	OK
Net II	Lewiston, Maine	1/91-	none	OK
	Upper Dam, Maine	1/91-	none	

TABLE 7.—Continued

Net	Station	Dates used	Recorded moves	Double-mass check
Net I	Belvidere, N.J.	9/93-	none	OK
	Boonton, N.J.	1/07-	none	
	New Brunswick, N.J.	1/96-	none	
Net II	*Woodstock College, Md.	1/80-	? , 1/14	1 break, adjusted before 1914
	Dale Enterprise, Va.	1/80-	none	
	Fallston, Md.	1/80-4/53	none	
	Rowlesburg, W. Va.	1/85-	8/26	
Net I	Philo (35W), Ohio	10/07-	none	
	Philo (#2), Ohio	1/02-9/65	none	
Net I	Tarboro, N.C.	1/00-	1/61, 3/66	Adjusted after 1/61-winter only
Net II	Sloan, N.C.	1/93-	none	2 breaks, 11/15& 5/62
	Kinston, N.C.	9/99-	6/36, 3/50, 21/52, 8/64	
	Edenton, N.C.	1/96-	none	
	Rocky Mount, N.C.	1/05-	none	
Net I	West Point, Ga.	1/99-	2/65	OK
Net II	Rome, Ga.	1/70-	7/53, 2/62	OK
	Lewisburg (Exp. Farm), Tenn.	1/99-	10/38	
	Palmetto, Tenn.	7/05-	none	
Net I	*St. Leo, Fla.	10/07	7/64	Not done

* Denotes ESSA-USWB climatological reference station. Unless otherwise indicated, last date used is 12/66.

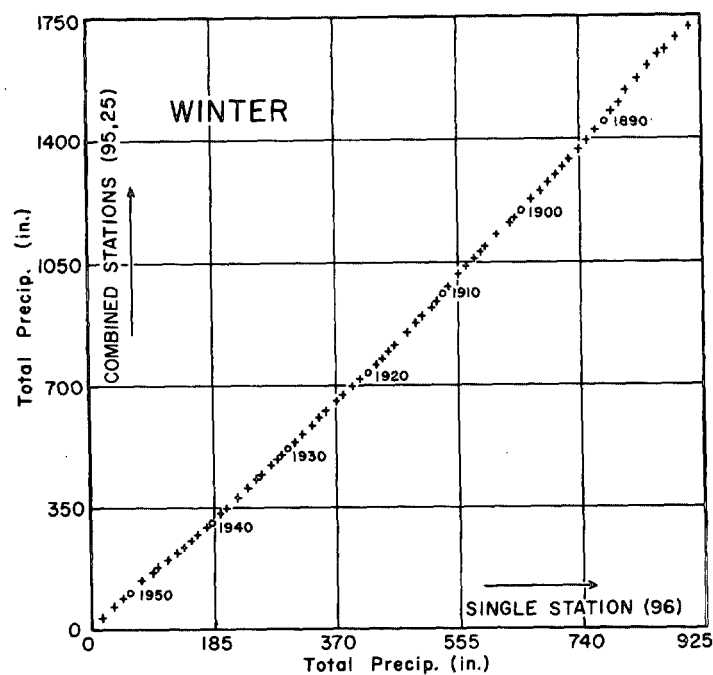


FIGURE 9.—Double-mass plot for winter precipitation, Woodstock College, Md. (station #96), versus the sum of Fallston, Md. (station #95), and Dale Enterprise, Va. (station #25).

The importance of the cooperative weather station in the production of homogeneous records that can be used in studies of climatic variations is underscored by the entries in table 7. Some of these stations have been maintained by institutions or by generations of a family in the same location for 50 or more years. Without these records, studies such as the present one would be impossible—I would like to take this opportunity to commend these institutions and people.

REFERENCES

- Brier, Glenn W., "Some New Approaches to the Non-Linear Prediction of Climatological and Hydrological Time Series," *Proceedings of the Conference on the Drought in the Northeastern United States, New York University, May 15-17, 1967, Geophysical Research Laboratory TR-68-3*, New York University, Mar. 1968, pp. 160-181.
- Bruce, J. P., and Clark, R. H., *Introduction to Hydrometeorology*, Pergamon Press, New York, 1966, 319 pp.
- Edwards, A., *Statistical Methods*, 2d Edition, Holt, Rinehart, and Winston, New York, 1967, 462 pp.
- Fiering, Myron B., "Multivariate Technique for Synthetic Hydrology," *Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers*, Vol. 90, No. 5, Sept. 1964, pp. 43-60.
- Fiering, Myron B., *Streamflow Synthesis*, Harvard University Press, Cambridge, Mass., 1967, 139 pp.
- Freiberger, W., and Grenander, U., "On the Formulation of Statistical Meteorology," *Review of International Statistical Institute*, Vol. 33, No. 1, Institut international de statistique, The Hague, 1965, pp. 59-86.
- Glejser, H., "A New Test for Heteroskedasticity," *Journal of the American Statistical Association*, Vol. 64, No. 325, Mar. 1969, pp. 316-323.
- Gray, James R., *Probability*, Interscience Publishers, Inc., New York, 1967, 267 pp.
- Grimmer, M., "The Space-Filtering of Monthly Surface Temperature Anomaly Data in Terms of Pattern, Using Empirical Orthogonal Functions," *Quarterly Journal of the Royal Meteorological Society*, Vol. 89, No. 381, July 1963, pp. 395-408.
- Holmström, Ingemar, "On a Method for Parametric Representation of the State of the Atmosphere," *Tellus*, Vol. 15, No. 2, May 1963, pp. 127-149.
- Hoyt, William G., and Langbein, Walter B., "The Yield of Streams as a Measure of Climatic Fluctuations," *Geographical Review*, Vol. 34, No. 2, Apr. 1944, pp. 218-234.
- Kendall, M., *Rank Correlation Methods*, Hafner Publishing Co., New York, 1962, 199 pp.
- Kendall, Maurice, and Stuart, Alan, *The Advanced Theory of Statistics*, Vol. 3, Hafner Publishing Co., New York, 1966, 552 pp.
- Kutzbach, J., "Empirical Eigenvectors of Sea-Level Pressure, Surface Temperature, and Precipitation Complexes Over North America," *Journal of Applied Meteorology*, Vol. 6, No. 5, Oct. 1967, pp. 791-802.
- Landsberg, H. E., Mitchell, J. M., Jr., and Crutcher, H. L., "Power Spectrum Analysis of Climatological Data for Woodstock College, Maryland," *Monthly Weather Review*, Vol. 87, No. 8, Aug. 1959, pp. 283-298.
- Langbein, Walter B., "Variations and Trends in River Flow," *Water Resources Review*, U.S. Geological Survey and Canada's Department of Energy, Mines, and Resources, Washington, D.C., May 1968, 10 pp.
- Langbein, Walter B., and others, "Annual Runoff in the United States," *U.S. Geological Survey Circular No. 52*, Washington, D.C., June 1949, 14 pp.
- Linsley, Ray K., Jr., Kohler, Max A., and Paulhus, Joseph L. H., *Applied Hydrology*, McGraw-Hill Book Co., Inc., New York, 1949, 689 pp.
- Lorenz, Edward N., "Empirical Orthogonal Functions and Statistical Weather Prediction," *Scientific Report No. 1, Contract No. AF19(604)1566*, Department of Meteorology, Massachusetts Institute of Technology, Cambridge, Dec. 1956, 49 pp.
- McDonald, James E., and Green, Christine R., "A Comparison of Rank-Difference and Product-Moment Correlation of Precipitation Data," *Journal of Geophysical Research*, Vol. 65, No. 1, Jan. 1960, pp. 333-336.
- Mitchell, J. Murray, Jr., Environmental Data Service, ESSA, Silver Spring, Md., 1960, (personal conversation).
- Mitchell, J. Murray, Jr., "The Measurement of Secular Temperature Change in the Eastern United States," *U.S. Weather Bureau Research Paper No. 43*, Washington, D.C., Mar. 1961, 80 pp.
- O'Brien, J. J., and Griffiths, J. F., "Choosing a Test of Normality for Small Samples," *Archiv für Meteorologie, Geophysik, und Bioklimatologie*, Ser. A, Vol. 16, No. 2/3, Vienna, 1967, pp. 267-272.
- Pearson, Egon S., and Hartley, H., *Biometrika Tables for Statisticians*, Vol. 1, Cambridge University Press, England, 1962, 240 pp.
- Sellers, William, "Climatology of Monthly Precipitation Patterns in the Western United States, 1931-1966," *Monthly Weather Review*, Vol. 96, No. 9, Sept. 1968, pp. 585-595.
- Stidd, C., "The Use of Eigenvectors for Climatic Estimates," *Journal of Applied Meteorology*, Vol. 6, No. 2, Apr. 1967, pp. 255-264.
- Subcommittee on Hydrology, Canadian Department of Energy, Mines, and Resources, "Statistical Methods in Hydrology," *Proceedings of Hydrology Symposium No. 5, Montreal, February 24-25, 1966*, Queen's Printer, Ottawa, 1967, 424 pp.
- U.S. Geological Survey, "General Introduction and Hydrologic Definitions," *Geological Survey Water-Supply Paper 1541-A*, Washington, D.C., 1960, 29 pp.
- U.S. Geological Survey, *Water Supply Papers*, U.S. Department of Interior, Washington, D.C., 1908-1966.
- U.S. Weather Bureau, ESSA, *Climatology of the United States—Climatic Summary of the United States (Bulletin W and Supplements)*, U.S. Department of Commerce, Washington, D.C., 1930, 1954, and 1960.
- U.S. Weather Bureau, "Substation History," *Key to Meteorological Records Documentation No. 1.1*, separate publication for each State, Washington, D.C., 1956.
- Williams, Philip, Jr., and Peck, Eugene L., "Terrain Influences on Precipitation in the Intermountain West as Related to Synoptic Situations," *Journal of Applied Meteorology*, Vol. 1, No. 3, Sept. 1962, pp. 343-347.
- Yevjevich, Vujica M., "Patterns in Sequence of Annual River Flow and Annual Precipitation," *Proceedings of the IUGG Symposium, Berkeley, Calif., 19-31 August 1963, Publication No. 63*, WMO/IUGG, Gentbrugge, Belgium, 1964, pp. 233-252.

[Received June 16, 1969; revised September 2, 1969]